

## CLAIMS

What is claimed is:

1. A method for enhancing voice activity detection comprising:  
determining a peak-to-mean likelihood ratio; and  
comparing the peak-to-mean likelihood ratio to a selected threshold to  
determine whether a current audio frame represents a voice signal.

2. The method of claim 1, wherein prior to determining the peak-to-mean likelihood ratio, the method further comprises:  
determining a short-term averaged energy for the current audio frame;  
and  
determining a long-term averaged energy for the current audio frame.

3. The method of claim 2, wherein after determining the short-term averaged energy and the long-term averaged energy, the method further comprises:  
determining whether a sum of the short-term averaged energy and a factor is greater than the long-term averaged energy; and  
determining that the current audio frame represents silence if the sum is less than the long-term averaged energy, without necessitating a determination of the peak-to-mean likelihood ratio.

Sub  
E1

1 4. The method of claim 3, upon determining that the sum is  
2 greater than the long-term averaged energy and before determining the peak-  
3 to-mean likelihood ratio, the method further comprises:  
4 determining whether a difference between the long-term averaged  
5 energy and the short-term averaged energy is less than a predetermined  
6 threshold;  
7 determining that the current audio frame represents voice if the  
8 difference is greater than the predetermined threshold; and  
9 continuing by determining the peak-to-mean likelihood ratio if the  
10 difference is less than the predetermined threshold.

1 5. The method of claim 2, wherein the determining of the short-  
2 term averaged energy comprises:  
3 determining an energy, in decibels, of the current audio frame;  
4 determining a short-term averaged energy for a prior audio frame; and  
5 conducting a weighted average of the energy of the current audio frame  
6 and the short-term averaged energy for the prior audio frame.

1 6. The method of claim 1, wherein the determining a peak-to-  
2 mean likelihood ratio comprises  
3 calculating an averaged peak-to-mean ratio for the current audio  
4 frame;  
5 determining a maximum averaged peak-to-mean ratio;  
6 determining a minimum averaged peak-to-mean ratio;

7 determining a first result being a difference between the maximum  
8 averaged peak-to-mean ratio and the averaged peak-to-mean ratio for the  
9 current audio frame;  
10 determining a second result being a difference between the maximum  
11 averaged peak-to-mean ratio and the minimum averaged peak-to-mean ratio;  
12 and  
13 conducting a ratio between the first result and the second result to  
14 produce the peak-to-mean likelihood ratio.

1 7. A communication module comprising:  
2 a substrate;  
3 a processing unit placed on the substrate; and  
4 a memory coupled to the processing unit, the memory to contain a  
5 voice activity detector which, when executed by the processing unit, analyzes  
6 a short-term averaged energy, a long-term averaged energy, and a peak-to-  
7 mean likelihood ratio in order to determine whether a current audio frame  
8 represents voice or silence.

56  
C3  
1 8. The communication module of claim 7, wherein the voice  
2 activity detector, when executed, controls the processing unit to determine  
3 whether a sum of the short-term averaged energy and a predetermined factor  
4 is greater than the long-term averaged energy, and to signal that the current  
5 audio frame represents silence if the sum is less than the long-term averaged  
6 energy.

Sub  
E3  
1 9. The communication module of claim 8, wherein the voice  
2 activity detector, when executed, controls the processing unit to determine

Sub  
F3  
3 whether a difference between the long-term averaged energy and the short-  
4 term averaged energy is less than a predetermined threshold, and to signal  
5 that the current audio frame represents voice if the difference is greater than  
6 the predetermined threshold.

1 10. The communication module of claim 9, wherein the voice  
2 activity detector, when executed, controls the processing unit to determine  
3 the peak-to-mean likelihood ratio, and to compare the peak-to-mean  
4 likelihood ratio to a selected threshold to determine whether a current audio  
5 frame represents a voice signal.

1 11. The communication module of claim 10, wherein the voice  
2 activity detector, when executed, controls the processing unit to determine a  
3 peak-to-mean ratio by (i) sampling an analog signal a predetermined number  
4 of times to produce a plurality of sampled signals each having a sampled  
5 value, (ii) determining a maximum value of the plurality of sampled signals,  
6 and (iii) conducting a ratio between an absolute value of the maximum value  
7 and a summation of the sampled values for the plurality of sampled signals.

Sub  
C4  
1 12. The communication module of claim 10, wherein the voice  
2 activity detector, when executed, controls the processing unit to determine an  
3 averaged peak-to-mean ratio for the current audio frame by (i) monitoring a  
4 maximum averaged peak-to-mean ratio and a minimum averaged peak-to-  
5 mean ratio, (ii) determining a first result being a difference between the  
6 maximum averaged peak-to-mean ratio and the averaged peak-to-mean ratio  
7 for the current audio frame, (iii) determining a second result being a  
8 difference between the maximum averaged peak-to-mean ratio and the

9 minimum averaged peak-to-mean ratio, and (iv) conducting a ratio between  
10 the first result and the second result to produce the peak-to-mean likelihood  
11 ratio.

13. A machine readable medium having embodied thereon a  
computer program for processing by a machine, the computer program  
comprising:

a first routine for determining a peak-to-mean likelihood ratio; and  
a second routine for comparing the peak-to-mean likelihood ratio to a  
selected threshold to determine whether an audio frame being transmitted  
represents a voice signal.

14. The machine readable medium of claim 13, wherein the  
computer program further comprising:

a third routine for determining a short-term averaged energy for the  
audio frame, the third routine being executed before the first and second  
routines; and

a fourth routine for determining a long-term averaged energy for the  
audio frame, the fourth routine being executed before the first and second  
routines.

15. The machine readable medium of claim 14, wherein the  
computer program further comprising:

a fifth routine for determining whether a sum of the short-term  
averaged energy and a predetermined factor is greater than the long-term  
averaged energy, the fifth routine being executed before the first and second  
routines; and

Sub  
E6

7 a sixth routine for determining whether a difference between the long-  
8 term averaged energy and the short-term averaged energy is less than a  
9 predetermined threshold, the sixth routine being executed after determining  
10 that the sum is greater than the long-term averaged energy and before  
11 execution of the first and second routines.

Sub  
X4

1 16. The machine readable medium of claim 15, wherein the fifth  
2 routine determining that the current audio frame represents silence if the  
3 sum is less than the long-term averaged energy.

1 17. The machine readable medium of claim 15, wherein the sixth  
2 routine determining that the current audio frame represents voice if the  
3 difference is greater than the predetermined threshold.

00134272-081498

Sub  
X5

1 18. A voice activity detector comprising:  
2 circuitry to determine a short-term averaged energy for an audio frame;  
3 circuitry to determine a long-term averaged energy for the audio frame;  
4 circuitry to determine whether the short-term averaged energy is  
5 greater than the long-term averaged energy by a predetermined factor;  
6 circuitry to determine whether a difference between the long-term  
7 averaged energy and the short-term averaged energy is less than a  
8 predetermined threshold when the short-term averaged energy is greater  
9 than the long-term averaged energy by the predetermined factor;  
10 circuitry to determine a peak-to-mean likelihood ratio when the  
11 difference between the long-term averaged energy and the short-term  
12 averaged energy is less than the predetermined threshold; and

- 13           circuitry to comparing the peak-to-mean likelihood ratio to a selected  
14   threshold and to determine that the audio frame represents a voice signal  
15   when the peak-to-mean likelihood ratio is greater than a selected threshold.

add  
A1

add  
C8

Add  
E9

004490-24ET60